МАТЕМАТИЧЕСКИЕ МЕТОДЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ГУМАНИТАРНЫХ ИССЛЕДОВАНИЯХ

STATISTICAL ANALYSIS OF INDIVIDUAL TASKS ON PROBABILITY THEORY

K. A. Kostenko, Y. Y. Katsman
(Tomsk, Tomsk Polytechnic University)
Tomsk, Russia
kak31@tpu.ru

Abstract: Over the years, improving the quality of basic education programs (PLO) in Tomsk Polytechnic University has received increased attention. One of the main objectives of improving the educational process and the PLO is the optimization of procedures for monitoring the quality of the PLO for their continuous improvement. This work contains the results of the statistical analysis of the quality of the tests for monitoring students' probability theory knowledge. The analysis showed a significant difference (not parallel) ofvariants of individual tasks, and on its basis is method for providing parallel tests.

Keywords: statistical analysis; probability theory; testing; scatter plot; sample characteristic; rank; median; cluster

I. Introduction. Today assessment of the quality of test materials that are used to test the knowledge and skills of students is quite time-consuming and difficult task, which is relevant for new disciplines, and for those training in production for several years [1]. Test materials are usually presented in the several variants, so that there is a problem of parallelism that can complicate the assessment of students' knowledge and its objectivity.

Usually the analysis of quality control materials much attention is paid to the parallel variants of the task [2, 3]. However, if the use of the modern theory test – Item Response Theory (IRT) [4] for estimate the latent factors required to provide for one test a minimum sample size of 200 to 1000 observations, the classical statistical theory allows us to obtain the estimates of the parameters, limited to a much smaller number of experiments.

II. The main aim of the study. The main aim of this work was a statistical analysis of parallel variants of the individual tasks on probability theory for assesses the quality of knowledge learned by students.

III. The formulation of the task. On the results of the control work (testing) the minimum score (3 points) was given for the attempt to solve at least one task, the maximum score (15 points) – the right solution for three tasks. Statistical analysis of the all the results of the control work on probability theory were processed in licensed program Statistica. Furthermore, it was required to determine the equivalence of the variants through using Kruskal-Wallis ANOVA test, Median test and Sheffe test.

IV. Analysis. Before analyzing the parallel of the variants was used the module of the descriptive statistics and was excluded unrepresentative variants (4 and under observation).

Further, it was assumed that these options are parallel (equivalent), and then the evaluation of the students should be adequate to their knowledge, rather than the complexity of tickets. Therefore for each option were calculated point and interval estimates that given the random factors assumed approximately equal average scores and variances for each variant. Realistic estimates for each variant are shown in Fig. 1 as a scatter plot.

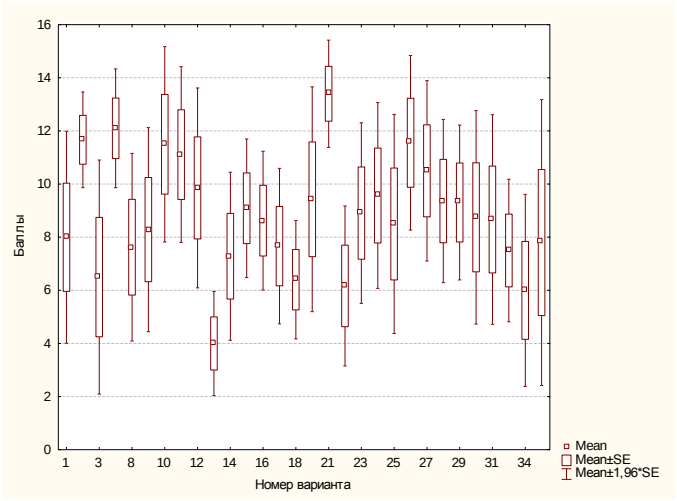

Fig 1.Scatterplots for different variants of the test

The present results clearly showed uneven difficulty (not parallel) of different variants of test. Partitioning variants for complexity group was conducted by the method of cluster analysis, k-means and only one variable – points, so to ensure at-dimensional equal number of cases in each group and the homogeneity of observations within the group all variants sorted by the average score:

– with complex tasks, where the average score was less than 7.8;

– with the tasks of medium complexity, the average score which ranged between 7.8 and 9.3;

– with simple jobs, which exceeded the average score of 9.3.

The module descriptive statistics instrumental in obtaining point estimates for all observations and each cluster separately. With this module the following results:

– the maximum difference in the estimates for the first and third cluster is less than three points;

– almost all point features for the second cluster and the entire set of observations are the same;

– for all cases 50% of the second cluster of the results exceeded 8.6 points, at the same time for the first cluster 50% of the results did not exceed 6.6 scores, and for the third cluster count exceeded 50% 10.5 points;

– variances for all observations and clusters were considered almost equal (the ratio of the variances of less than 2);

– the analysis of the factors skew and kurtosis testified that the distribution of scores in each group are asymmetric and differs substantially from the Gaussian distribution.

For test the hypothesis of a significant influence of the factorswas conducted Oneway analysis package Statistica with the marked tests are significant at $p <,05000$. Analysis of the sum of rank in groups (clusters) obtained as a result of the Kruskal-Wallis test [5]confirmed that the maximum score was observed in a cluster with the options had an easy task, and the minimum was observed in the cluster with the options had difficult tasks.

Analysis of the results of another type of rank test – the median test, presented in a table shows that:

– the top half of the table contains the maximum value of the cluster, which corresponded to tasks with a high level of complexity and they produce minimum estimates;

– the bottom half of the table included a maximum value of clusters, which corresponded to tasks with a low level of difficulty and they produce maximum estimates.

The hypothesis of the influence of factors, tested during the median test, as well as the analysis of Kruskal-Wallis test showed that the effect of a significant factor.

Further by analyzing the Mann-Whitney hypothesis was tested at two different sample homogeneity (clusters)[5].

Evaluate the effects of processing are shown graphically in Fig. 2.

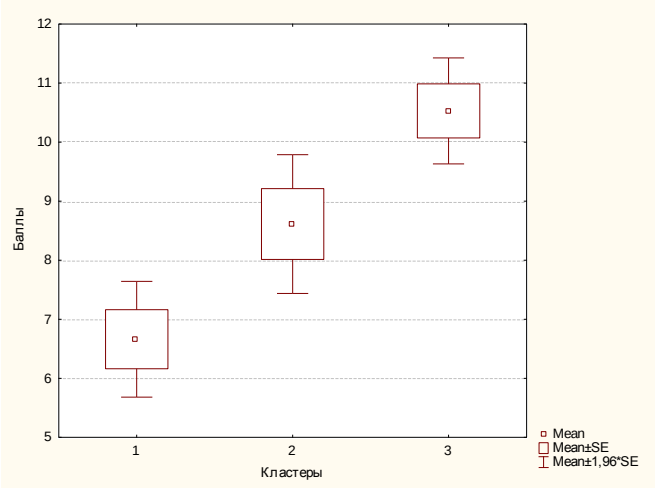

Fig 2.Scatterplots for all clusters

The results in Fig. 2 show a significant difference in the characteristics of point and interval for different groups. By analyzing them it was possible to try to answer the question: what a pair of groups of tasks variants can be considered significantly different? To answer this question we compared the average Sheffe test's method for different pairs of levels of the factors, the analysis of which showed a significant difference between the average scores for various pairs of clusters that proves the validity of the alternative hypothesis of a slight impact factor [5].

V. Conclusion. Monitoring the quality of teaching subjects to a large extent determined by the quality control of individual teaching materials (tests). One of the major characteristics of the options is to test them in parallel.

Statistical analysis of the monitoring of individual tasks on probability theory showed that variants of test tasks are not parallel. On the basis of the research was to draw definitive conclusions about the quality of the proposed test items: 4 of 39 the available options for individual tasks were excluded due to non-representative sampling, 9 variants contain the problem complex level with minimal received them estimates, 9 variants with the objectives of mid-level and 12 options with the objectives of easy level.

Conducted research have shown that the first and second clusters can be expanded by adding new options to exclude from the third cluster versions with two or three easy task, considering the average of points. And replace them with more complex tasks of the first and second clusters, also considering the dialed the average of points.

This paper shows that even for tasks that are used for a number of years, the task of ensuring the parallelism is relevant. The proposed in the work statistical methods allowsuccessfully solve this problem as demonstrated by the example of the control tasks of the theory of probability.

References

[1] Y.Y. Katsman "Statistical analysis of individual tasks on probability theory". In: Bulletin of the Tomsk Polytechnic University (Bulletin TPU), Tomsk, Russia, 2014, pp. 84-90. (style for Paper in a Proceedings).

[2] H.K. Suen, P.W. Lei “Classical versus Generalizability theory of measurement”. Available at: http://suen.educ.psu.edu/~hsuen/pubs/Gtheory.pdf (accessed on 10.03.2016).
[3] B.V.Ilyukhin, O.E. Permyakov “Quality assurance issues and directions in perfecting the system of competitive selection of applicants in the Russian Federation”.Tomsk Polytechnic University, Tomsk, Russia, 2007. (style for Book).
[4] G.Rasch “Probabilistic models for some intelligence and attainment tests”. Copenhagen: Danish Institute for Educational Research, Copenhagen, Denmark, 1960. (style for Book).
[5] Ya.Gaek, E.Shidak “Теория рангов и их критериев” [Theory of rank tests]. Moscow, Nauka Publ., 1971, 374 p.

ПРОГНОЗ НЕБЛАГОПРИЯТНОГО ИСХОДА ЛЕЧЕНИЯ НА ОСНОВЕ КОМПЛЕКСНОГО ОЦЕНИВАНИЯ ФАКТОРОВ РИСКА

Аброськина А.А.
(г. Кемерово, Кемеровский государственный университет)
abroskina.anastasiya@mail.ru

PROGNOSIS IS POOR TREATMENT OUTCOME BASED ON A COMPREHENSIVE ASSESSMENT OF THE FACTORS RISK

Abros’kina A.A.
(Kemerovo, Kemerovo State University)

Annotation. The article discusses different approaches to building a comprehensive risk assessment for the prediction of adverse outcome of treatment.

Keywords. Risk factors, a comprehensive approach, logistic regression, the integral index, comprehensive risk assessment.

Одной из наиболее важных задач для медицинского работника при работе с пациентом является задача прогноза исхода проводимого лечебного мероприятия. Поэтому в настоящее время наиболее важная роль принадлежит индивидуальному подходу к больному. Повысить эффективность проводимых мероприятий, можно определив наиболее значимые факторы риска и оценку их влияния на лечение больного.

В настоящее время для выбора факторов, предположительно влияющих на исход лечебного мероприятия, широко применяется бинарная логистическая регрессия. Однако, при её построении может возникнуть проблема, заключающаяся в том, что все факторы риска измерены в качественных шкалах: номинальной, ранговой. С математической точки зрения построение бинарной логистической регрессии только по качественным данным является некорректным.

Один из подходов решения данной проблемы заключается в комплексном оценивании исследуемого процесса, то есть в представлении всей группы показателей в виде одного интегрального показателя. Интегральный показатель, характеризующий комплексную оценку факторов риска, позволяет решать задачу снижения размерности с минимальной потерей информации.

В медицинских исследованиях существуют различные подходы к комплексному оцениванию факторов риска. Чаще всего комплексная оценка процесса, представляет собой сумму баллов, выставленных разным уровням факторов риска. Например, пациент имеет сахарный диабет – 1 балл, не имеет – 0 баллов; принимает диуретики – 1 балл, нет – 0 баллов, острая сердечная недостаточность – 1 балл, нет – 0 баллов. Построенный таким образом